

Санкт-Петербургский политехнический университет  
Петра Великого  
Физико-механический институт  
Высшая школа прикладной математики и вычислительной физики

Реферат по дисциплине  
"Автоматизация научных исследований"

**Автоматическая генерация научной работы с помощью LLM**

Выполнил студент:

Бабахина Софья Александровна

Группа: 5040102/50201

Санкт-Петербург  
2025

# Содержание

<b>1</b>	<b>Введение</b>	<b>2</b>
<b>2</b>	<b>Описание экспериментов</b>	<b>2</b>
2.1	Исходные данные и ожидаемые результаты . . . . .	2
2.2	Постановка задач для LLM . . . . .	2
2.3	Результаты генерации . . . . .	3
2.3.1	Генерация источников . . . . .	3
2.3.2	Генерация наброска текста работы . . . . .	4
2.3.3	Генерация UML по описанию . . . . .	4
2.3.4	Генерация листинга для веб-сервиса . . . . .	6
2.3.5	Генерация аннотации . . . . .	8
2.3.6	Перевод аннотации на английский . . . . .	10
2.3.7	Поиск списка литературы: генерация списка ключевых слов . . . . .	11
2.3.8	Проверка на соответствие структуре научного текста . . . . .	13
<b>3</b>	<b>Заключение</b>	<b>14</b>

# 1 Введение

**Цель работы:** Проверить, можно ли, используя исключительно LLM, написать научную работу на заданную тему.

**Задание:** Сгенерировать научную работу с помощью LLM и подготовить целостный отчет на тему «Как автоматизировать написание научной работы» с описанием и анализом использованных при генерации промптов из проделанных лабораторных и примеров их работы на основе сгенерированного с помощью ИИ текста научной работы.

## 2 Описание экспериментов

### 2.1 Исходные данные и ожидаемые результаты

В качестве исходных данных для работы с LLM использовался текст научно-исследовательской работы, посвящённой автоматическому детектированию пилообразных колебаний плазмы по данным диагностики мягкого рентгеновского излучения на токамаке «Глобус-М2», **который также был сгенерирован с помощью ИИ**. В качестве «сырых» данных для генерации были использованы название будущей работы и описанные в свободной форме планы и задачи на нир/статью:

- **Название:** «Автоматическое детектирование пилообразных колебаний по данным SXR-диагностики на токамаке «Глобус-М2»».
- **Короткое описание запланированных работ:** Эта тема находится на стыке физики плазмы, анализа данных и машинного обучения. В этой работе я планирую провести исследование, написать программу для детектирования и собрать дистрибутив/сайт в локальной сети, которым смогут пользоваться сотрудники моей лаборатории. Планируется, что программа будет использовать для автоматического детектирования методы машинного обучения, математической статистики и математического анализа.

На выходе ожидается ряд «продуктов» ИИ:

- Файл нир/статьи/шаблона магистерской работы, собранный по частям в ходе общения с разными моделями ИИ.
- Файл с метаданными (для статьи, например), содержащий аннотацию, ключевые слова и список литературы в переводе на английский язык.
- Набор артефактов - протоколов общения с ИИ (DeepSeek, chatGPT и проч.): все основные диалоги в формате длинных скриншотов, результаты генераций (файлы для листинга) и т.д.
- Настоящий реферат, являющийся, фактически, отчетом по проделанной работе.

### 2.2 Постановка задач для LLM

Задачи для LLM можно разделить на два этапа:

- Генерация текста для последующего улучшения и анализа.
- Проведение анализа сгенерированного на первом этапе научного текста и формирование вспомогательных элементов, обеспечивающих его соответствие стандартам научной работы.

Первый этап исследования включает в себя анализ литературы без формального участия ИИ, что необходимо для отбора наиболее подходящих источников и генерации наброска текста работы на их основе. На этом этапе ИИ используется для поиска источников, относящихся к предметной области, по названию работы и короткому описанию (Лаб.6), а также для генерации шаблона текста нир/статьи. Во второй части исследования можно выделить ряд более конкретных задач:

- Генерация диаграммы UML, описывающей предметную область или способы взаимодействия с созданным программным обеспечением/сервисом/программой. - Лаб.3
- Генерация листинга сайта, обеспечивающего взаимодействие с разработанным программным обеспечением/сервисом/программой для потенциальных пользователей (работников лаборатории). - Лаб.5
- Генерация аннотации к готовому тексту научной работы. - Лаб.1
- Генерация перевода аннотации из предыдущего пункта. - Лаб.2
- Генерация запроса в научный архив arXiv (<https://arxiv.org>) для пополнения списка литературы актуальными источниками из предметной области исследования. - Лаб.7
- Анализ сгенерированного текста и выявление наличия или отсутствия обязательных структурных элементов научной работы (актуальность, цель, задачи, объект, предмет, методология, результаты, выводы). - Лаб.4

*В следующем разделе каждая из задач будет рассмотрена подробнее.*

## 2.3 Результаты генерации

### 2.3.1 Генерация источников

Для генерации источников, описывающих предметную область, необходимых для предварительного ознакомления с темой работы, были использованы две модели: chatGPT и DeepSeek. За основу было взято исследование, проведенное в 6 лабораторной.

Текст промпта:

Ты будешь выступать в роли специалиста - математика, работающего в области физики плазмы. У меня есть примерная формулировка темы моей будущей магистерской работы: «Автоматическое детектирование пилообразных колебаний по данным SXR-диагностики на токамаке «Глобус-М2»». Короткое описание запланированных работ: Эта тема находится на стыке физики плазмы, анализа данных и машинного обучения. В этой работе я планирую написать программу для детектирования и собрать дистрибутив/приложение в докере/сайт в локальной сети, которым смогут пользоваться сотрудники моей лаборатории. Планируется, что программа будет использовать для автоматического детектирования методы машинного обучения, математической статистики и математического анализа. Чтобы можно было написать текст самой статьи, нужно составить примерный список литературы по данной теме. Используя предоставленное название и краткое описание, составь список релевантных научных источников. Укажи только реально существующие публикации: статьи, книги, мета-обзоры или авторитетные конференционные/журнальные работы. Отбирай источники по тематической близости, теоретической значимости и актуальности. Укажи валидные URL. Учитывай индекс цитируемости

Несмотря на результаты, полученные в ходе лабораторной работы 6, DeepSeek показал неутешительный результат: ни одна ссылка не указывала на валидный источник и лишь пару статей из списка удалось найти по названию через поисковую строку в браузере. В связи с чем было принято решение обратиться к chatGPT. Для проверки «состоятельности» предложенного списка литературы, помимо валидности ссылок, я проверяла наличие среди источников хотя бы одной из «опорных» статей, которыми мне предоставил научный руководитель. В списке, предложенном чатом GPT нашлось целых два источника из списка «опорных», поэтому было принято решение остановиться на этом варианте и не использовать другие модели.

### 2.3.2 Генерация наброска текста работы

Для генерации текста работы была использована модель DeepSeek. Выбор модели основывается на невозможности прикрепить в бесплатной версии других ИИ-моделей большого количества файлов для анализа. В DeepSeek было загружено 5 источников из списка, сгенерированного через chatGPT.

Текст промпта:

Ты будешь выступать в роли специалиста - математика, работающего в области физики плазмы. У меня есть примерная формулировка темы моей будущей магистерской работы: «Автоматическое детектирование пилообразных колебаний по данным SXR-диагностики на токамаке «Глобус-М2»». Короткое описание запланированных работ: Эта тема находится на стыке физики плазмы, анализа данных и машинного обучения. В этой работе я планирую провести исследование, написать программу для детектирования и собрать дистрибутив/сайт в локальной сети, которым смогут пользоваться сотрудники моей лаборатории. Планируется, что программа будет использовать для автоматического детектирования методы машинного обучения, математической статистики и математического анализа. Я подобрала несколько подходящих мне источников, которые прикрепляю во вложениях. Проанализируй их и выполни следующие требования:

1. Напиши шаблон-набросок для статьи по теме магистерской работы. Во вложении я прикрепила несколько источников, на которые ты можешь опираться при составлении текста.
2. Обязательно напиши вступление: опиши в нем актуальность и цели работы.
3. Обязательно напиши небольшой обзор литературы
4. Пиши связным текстом с выделениями по абзацам. Старайся не использовать списки и смайлики. Текст должен иметь научный стиль.

Полученный текст содержит описание актуальности задачи, обзор физических предпосылок возникновения пилообразных колебаний, анализ существующих подходов к их детектированию, формальную методологию предлагаемого алгоритма, результаты численных экспериментов на реальных диагностических данных, а также обсуждение практической применимости разработанного программного решения в экспериментальной среде токамака. Также были учтены требования к оформлению научных публикаций и отчетов научно-исследовательских работ, включая структурную полноту текста, логическую связность элементов исследования и корректность формулировок цели, методов и результатов.

*С полученным текстом можно ознакомиться в файле Шаблон.docx и протоколе общения с DeepSeek - DeepSeek-текст+листинг.pdf*

### 2.3.3 Генерация UML по описанию

С генерации UML диаграммы вариантов использования начинается второй этап исследования. Сгенерированный на первом этапе текст был сохранен в файл и загружен в chatGPT. За основу было взято исследование, проведенное в 3 лабораторной.

Текст промпта:

Я проанализировала список источников и набросала шаблон будущей работы. Проанализируй его и предложи вариант диаграммы вариантов использования будущей программы. Сгенерируй PlantUML код для диаграммы вариантов использования будущего дистрибутива/сервиса в локальной сети. Укажи действующие лица и перечисли случаи использования (дополни список вариантов сам) по ролям. Потенциальный набор действующих лиц такой:

Физик-экспериментатор: может использовать ПО в ходе экспериментов в реальном времени и т.д.

Физик-теоретик: может использовать на уже готовых данных для изучения природы пилообразных колебаний и т.д.

Разработчик (Я): разрабатываю ПО, дополняю, развиваю математический аппарат и т.д.

Инженер: Может использовать и реальном времени и на готовых данных, (что-то еще?)

```

1 @startuml
2 left to right direction
3 skinparam style strictuml
4 skinparam packageStyle rectangle
5 skinparam shadowing false
6
7 actor "Физик-\нэкспериментатор" as Exp
8 actor "Физик-\нтеоретик" as Theo
9 actor "Инженер" as Eng
10 actor "Разработчик" as Dev
11
12 rectangle "Сервис автоматического детектирования\нпилообразны колебаний\n
    (SXR, токамак Глобус-М2 )" {
13
14     (Загрузка данны \nSXR-диагностики) as UC_Load
15     (Предобработка\nSXR-сигналов) as UC_Pre
16     (Автоматическое\ндетектирование\nsawtooth-крашей) as UC_Detect
17     (Статистическая\nверификация\nсобытий) as UC_Stat
18     (Многоканальный\nпространственный\nанализ) as UC_Spatial
19
20     (Визуализация\nSXR-сигналов\nи детектированны \нкрашей) as UC_Vis
21     (Анализ параметров\nпилообразны \нколебаний) as UC_Params
22     (Экспорт результатов\nв csv, json,\нграфики) as UC_Export
23
24     (Пакетная обработка\nархивны данны ) as UC_Offline
25     (Работа в режиме\nквази-реального\nвремени) as UC_RT
26
27     (Контроль качества\nв одны данны \ндиагностики) as UC_QC
28     (Мониторинг\nработоспособности\nсервиса) as UC_Monitor
29
30     (Настройка параметров\nалгоритмов) as UC_Config
31     (Обучение и обновление\nML-моделей) as UC_ML
32 }
33
34 ' --- Связи акторов ---
35
36 Exp --> UC_Load
37 Exp --> UC_RT
38 Exp --> UC_Vis
39
40 Theo --> UC_Load
41 Theo --> UC_Offline
42 Theo --> UC_Vis
43 Theo --> UC_Params
44 Theo --> UC_Export
45
46 Eng --> UC_Load
47 Eng --> UC_RT
48 Eng --> UC_QC
49 Eng --> UC_Monitor
50 Eng --> UC_Vis
51
52 Dev --> UC_Config
53 Dev --> UC_ML
54 Dev --> UC_QC
55 Dev --> UC_Monitor
56
57 ' --- Включения (include) ---
58

```

```

59 UC_Detect <.. UC_Pre : <<include>>
60 UC_Detect <.. UC_Stat : <<include>>
61 UC_Detect <.. UC_Spatial : <<include>>
62
63 UC_Offline <.. UC_Detect : <<include>>
64 UC_RT <.. UC_Detect : <<include>>
65
66 UC_Params <.. UC_Detect : <<include>>
67 UC_Vis <.. UC_Detect : <<include>>
68
69 @enduml

```

Листинг 1: Код диаграммы вариантов использования (сгенерирован с помощью chatGPT)

Диаграмма читаема и по 10-балльной шкале соответствия UML-нотации может быть оценена на 8. Содержание и логичность могут быть оценены на 9 ввиду того, что четкого описания предоставлено не было. Результат отображения:

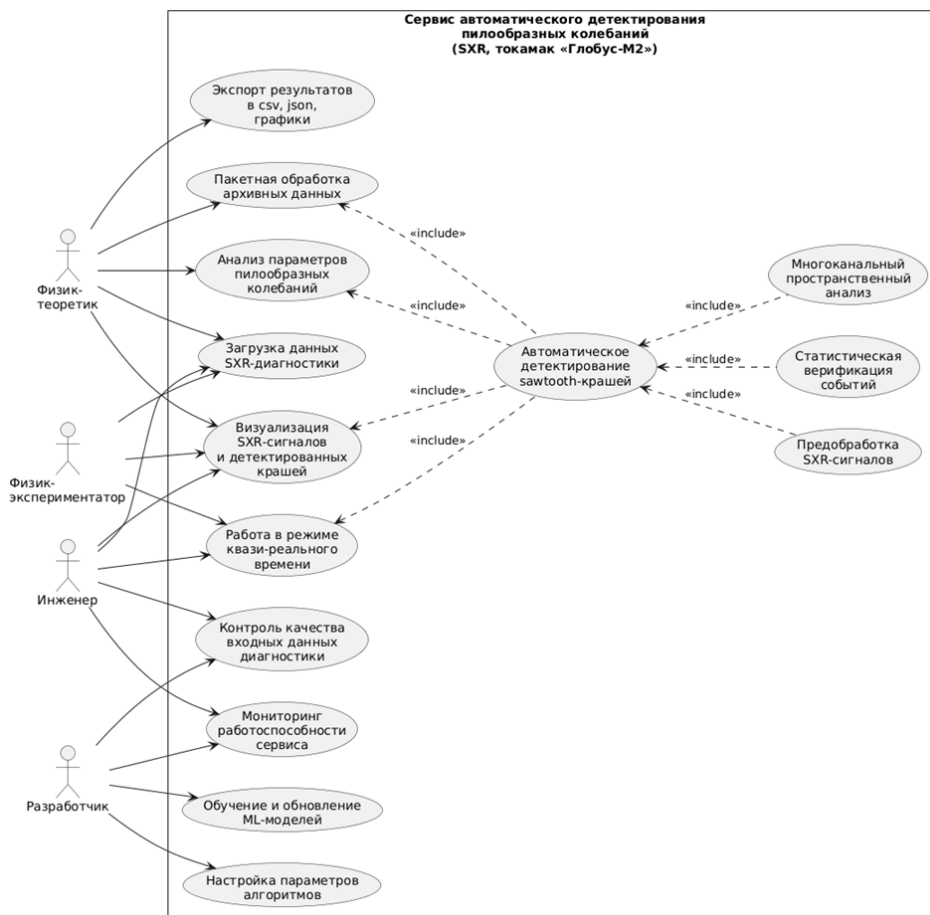


Рис. 1: Диаграмма вариантов использования

### 2.3.4 Генерация листинга для веб-сервиса

Для генерации листинга использовались две модели - DeepSeek и chatGPT. За основу было взято исследование, проведенное в 5 лабораторной. Сначала в продолжение диалога с GPT был отправлен следующий промпт: **Текст промпта для генерации:**

Проанализируй предоставленный текст на соответствие структуре и стандартам научной работы (статьи или введения к исследованию). Действуй строго по следующему алгоритму:

1. Идентификация структурных элементов. Найди в тексте и представь в виде таблицы с колонками: «Элемент», «Цитата или констатация отсутствия», «Краткий комментарий». Элементы для поиска:

Актуальность исследования, Цель исследования, Задачи исследования, Объект и предмет исследования, Методология / методы, Основные результаты, Выводы. Если элемент отсутствует или выражен неявно – четко отметить это в соответствующей ячейке.

## 2. Глубинный анализ и оценка.

На основе выявленных элементов проведи критический разбор:

Логическая связность: Проследи, вытекают ли задачи из цели, а выводы – из результатов.

Оцени, решает ли работа заявленную цель.

Корректность формулировок: Проверь, являются ли цели и задачи конкретными, измеримыми, а предмет и объект – четко разграниченными.

Выявление противоречий и лакун: Отметь смысловые или логические противоречия между разделами, а также структурные пробелы (например, отсутствие методологии).

## 3. Итоговый вердикт.

Представь заключение в трех блоках:

Структурная полнота: Общая оценка наличия всех обязательных элементов научного текста.

Ключевые замечания и ошибки: Список основных проблем, выявленных на этапе глубинного анализа.

Общий вывод: Резюме о соответствии текста структуре научной работы с рекомендациями по устранению главных недостатков.

Важно: Всё время опирайся только на предоставленный текст, не придумывай и не домысливай информацию.

Результат генерации полностью удовлетворил все требования, описанные в промпте, но не соответствовал моим ожиданиям. Исходя из результатов, полученным в ходе лабораторной работы номер 5, было принято решение обратиться с тем же промптом к DeepSeek. В итоге было принято решение остановиться на варианте, сгенерированном через DeepSeek, представляющем более широкий функционал и более приятный и удобный (большое окошко для вывода сигнала, возможность настройки расположения и проч.) интерфейса.

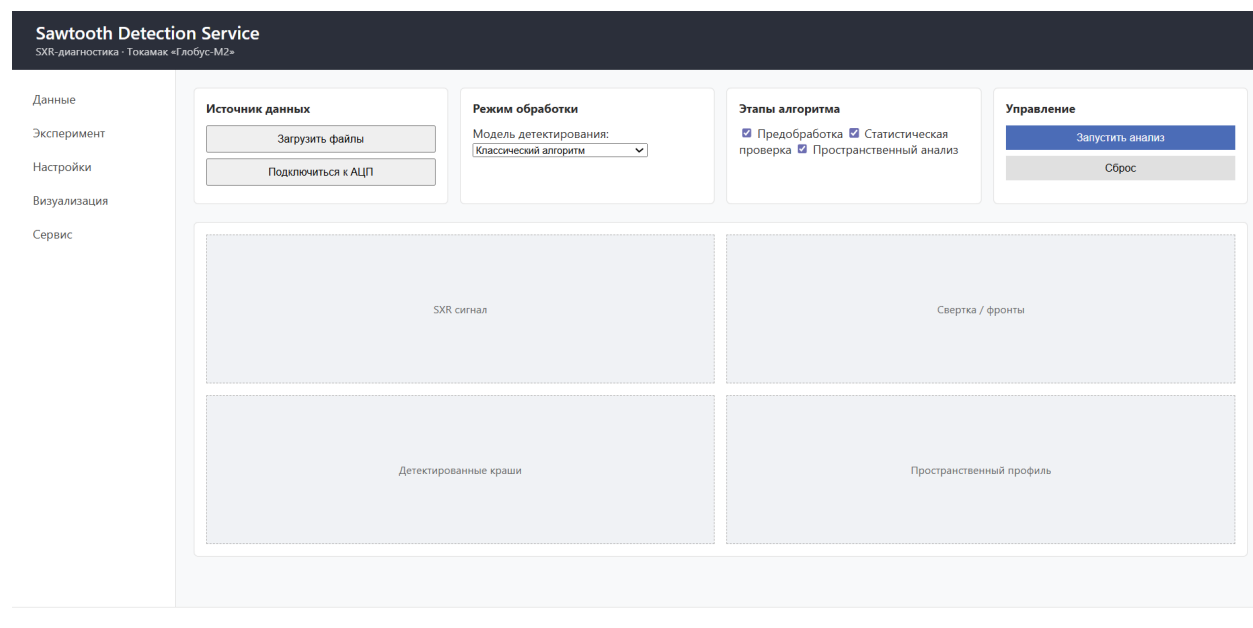


Рис. 2: Листинг GPT



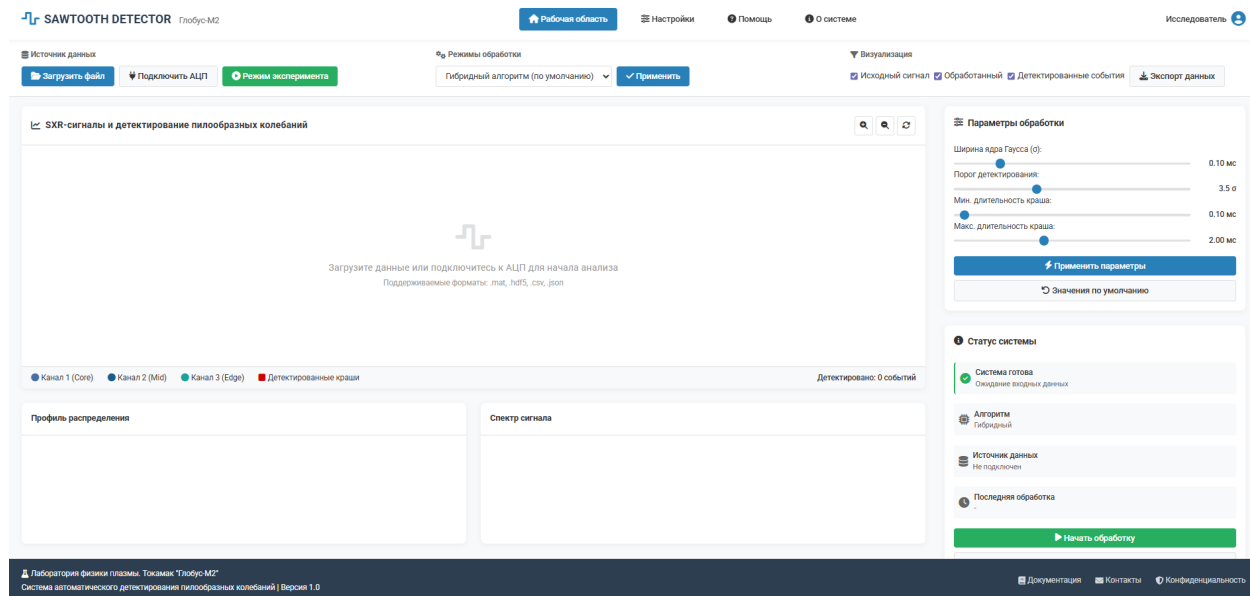


Рис. 3: Листинг DeepSeek

С кодом трех итоговых файлов, которые были оставлены в качестве «рабочего» варианта (с помощью DeepSeek), можно ознакомиться в приложенных к реферату файлах по расположению `source/listing/ds` (вынесены из текста ввиду большого объема). Код, сгенерированный в ходе беседы с chatGPT можно посмотреть в дополнительном файле - протоколе диалога с LLM (`chatGPT.png`), а также в самих файлах по расположению `source/listing/gpt`.

Результаты пп. 2.3.3 и 2.3.4 были размещены в блоке 4 научной работы «Реализация и программный дистрибутив» по рекомендации chatGPT.

### 2.3.5 Генерация аннотации

С помощью chatGPT (в том же чате) была сгенерирована аннотация. В основу было положено исследование из лабораторной работы 1.

Текст промпта:

Проанализируй загруженный текст научной работы и создай к нему аннотацию (файл Шаблон.docx). Аннотация – это один из видов сокращенной формы представления научного текста. Ее назначение – привлечь внимание читателя, пробудить читательский интерес сообщением сути исследования с помощью минимального количества языковых средств. Требования к структуре и содержанию аннотации:

Информативность и содержательность. Аннотация должна в обобщенном виде представлять содержание статьи. В тексте аннотации не стоит использовать общие фразы, а также указывать несущественные детали и общеизвестные положения. Предыстория (история вопроса) может быть приведена только в том случае, если она напрямую связана с раскрытием цели исследования.

Оригинальность. Следует избегать прямых повторов каких-либо фрагментов работы. Название статьи не должно дублироваться в тексте аннотации.

Четкость, логичность и связность изложения.

Компактность. Объем аннотаций к статьям должен составлять 50-100 слов.

В аннотации должны быть представлены следующие аспекты содержания статьи:

предмет, тема, цель работы (указываются в том случае, если они не ясны из заглавия статьи); метод или методология (их целесообразно описывать в том случае, если они отличаются новизной или представляют интерес с точки зрения данной работы); результаты работы (предпочтение отдается новым результатам и данным долгосрочного значения, важным открытиям, выводам, которые опровергают существующие теории, а также данным, которые, по мнению автора, имеют практическое значение);

выводы (они могут сопровождаться рекомендациями, оценками, предложениями, гипотезами, описанными в статье).

Не рекомендуется:

включать в аннотацию таблицы, рисунки, схемы, диаграммы и формулы;  
приводить ссылки на номера публикаций из пристатейного списка литературы;  
использовать сокращения и условные обозначения, кроме общеупотребительных. Если сокращения все же вводятся, то при первом употреблении необходимо дать их расшифровку.

Сгенерированный текст получился оптимальным по размеру, кроме того, ИИ-аннотация соответствует содержанию работы и стилю написания аннотаций к научным текстам. Результат генерации:

*В работе рассматривается задача автоматического детектирования пилообразных колебаний плазмы по данным диагностики мягкого рентгеновского излучения на токамаке «Глобус-М2». Целью исследования является разработка универсального и устойчивого алгоритма, способного надежно выявлять краши в широком диапазоне экспериментальных условий. Предложен гибридный метод, объединяющий анализ фронтов сигналов, адаптивную статистическую верификацию и многоканальный пространственный анализ. Алгоритм реализован в виде веб-сервиса и протестирован на архивных экспериментальных данных, продемонстрировав высокую точность и практическую применимость для оперативного анализа плазменных разрядов.*

Помимо аннотации в формате «уточняющего промпта» были запрошены ключевые слова на русском и английском языке.

Тексты промптов:

- **Исходный:** Сгенерируй список ключевых слов по текущей версии статьи. Нужны ключевые слова на русском, а также в переводе на английский
- **Уточняющий:** Сократи до 5 ключевых слов, сделай упор на машинное обучение и физику плазмы.

Результаты:

• **Ключевые слова (на русском)**

- пилообразные колебания;
- токамак;
- диагностика мягкого рентгеновского излучения;
- магнито-гидродинамические неустойчивости;
- автоматическое детектирование;
- обработка экспериментальных данных;
- анализ временных рядов;
- адаптивные статистические методы;
- многоканальный анализ;
- машинное обучение;
- веб-сервис для анализа данных;
- токамак «Глобус-М2».

**Keywords (in English)**

- sawtooth oscillations;
- tokamak;
- soft X-ray diagnostics;

- magnetohydrodynamic instabilities;
- automatic detection;
- experimental data processing;
- time series analysis;
- adaptive statistical methods;
- multi-channel analysis;
- machine learning;
- web-based data analysis service;
- Globus-M2 tokamak.

• **Ключевые слова (на русском)**

- пилообразные колебания;
- токамак;
- диагностика мягкого рентгеновского излучения;
- автоматическое детектирование;
- машинное обучение.

**Keywords (in English)**

- sawtooth oscillations;
- tokamak;
- soft X-ray diagnostics;
- automatic detection;
- machine learning.

*С результатами генерации аннотации и ключевых слов можно ознакомиться в протоколе беседы с ИИ [source/chatGPT.pdf](#)*

### 2.3.6 Перевод аннотации на английский

В отдельном чате, чтобы не назначать чату новую роль, был задан следующий промпт (за основу взято исследование из лабораторной работы 2):

Ты - эксперт-лингвист и носитель английского языка. Твоя задача - полностью самостоятельно проанализировать и перевести предоставленный русскоязычный текст. Проведи полный лингвистический анализ (определи тему, стиль, тон, ключевые слова и сложные места) и на его основе создай идеальный английский перевод. Перевод должен читаться как оригинальный текст, написанный специалистом в данной области. Все решения по выбору терминов и адаптации стиля принимай самостоятельно, исходя из контекста и лучших практик перевода. Текст для перевода: В работе рассматривается задача автоматического детектирования пилообразных колебаний плазмы по данным диагностики мягкого рентгеновского излучения на токамаке «Глобус-М2». Целью исследования является разработка универсального и устойчивого алгоритма, способного надежно выявлять краши в широком диапазоне экспериментальных условий. Предложен гибридный метод, объединяющий анализ фронтов сигналов, адаптивную статистическую верификацию и многоканальный пространственный анализ. Алгоритм реализован в виде веб-сервиса и протестирован на архивных экспериментальных данных, продемонстрировав высокую точность и практическую применимость для оперативного анализа плазменных разрядов.

Перевод соответствует русскому аналогу, терминология соблюдена, предложения не перегружены. Результат генерации:

*This work addresses the problem of automatic detection of plasma sawtooth oscillations using soft*

*X-ray diagnostic data from the Globus-M2 tokamak. The aim of the study is to develop a universal and robust algorithm capable of reliably identifying sawtooth crashes across a wide range of experimental conditions. A hybrid approach is proposed that combines signal edge analysis, adaptive statistical validation, and multichannel spatial analysis. The algorithm has been implemented as a web-based service and tested on archival experimental data, demonstrating high accuracy and practical applicability for the prompt analysis of plasma discharges.*

С результатами перевода аннотации на английский язык можно ознакомиться в протоколе беседы с ИИ [source/chatGPT-перевод аннотации.pdf](#)

### 2.3.7 Поиск списка литературы: генерация списка ключевых слов

Следующим шагом необходимо было наполнить исследование источниками. За основу было взято исследование из лабораторной работы 7. Была выбрана модель DeepSeek.

Текст промпта:

Сгенерируй для меня поисковые запросы для научного архива arXiv (<https://arxiv.org>).

Запросы должны быть рабочими и давать непустую выдачу. Следуй этим правилам:

1. Тематика и якорные термины: Все запросы должны быть связаны с одним из этих пяти обязательных "якорных" терминов. Используй каждый термин в нескольких запросах: sawtooth oscillations, tokamak, soft X-ray diagnostics, automatic detection, machine learning.

2. Формат и сложность: Создай 8 наборов запросов. Каждый набор должен содержать три уровня детализации, идущих друг за другом:

- Уровень 1 (Простой): 3-5 общих ключевых слов, разделённых пробелами. Не используй кавычки, операторы или специальные символы.

Пример: plasma diagnostics THOMSON SCATTERING

- Уровень 2 (С фразой): Одна короткая фраза из 2-3 слов в двойных кавычках, плюс 1-2 общих ключевых слова.

Пример: "Thomson scattering" plasma fusion

- Уровень 3 (Для расширенного поиска): Запрос, готовый для использования в расширенном интерфейсе или API arXiv, где можно указать поля. Используй конструкцию: abs:"краткая фраза" OR ti:"краткая фраза". Фраза должна быть до 3 слов.

Пример: abs:"Tukey boxplot" OR ti:"Tukey boxplot"

3. Общие ограничения:

- Используй только английский язык.
- Избегай узкоспециальных аббревиатур и очень длинных составных фраз.
- Не используй операторы AND, OR, NOT (кроме как в структуре Уровня 3), не указывай категории (cat:).
- Запросы должны быть практичными и реалистичными для поиска научных статей.

Выведи результат строго в следующем формате: для каждого из 8 наборов выведи три строки подряд - сначала Уровень 1, затем Уровень 2, затем Уровень 3. После каждого набора можно оставить пустую строку для читаемости.

Результат:

- sawtooth oscillations tokamak plasma instabilities  
"sawtooth oscillations" tokamak disruption  
abs:"sawtooth oscillation" OR ti:"sawtooth oscillation"
- tokamak fusion plasma confinement  
"tokamak plasma" magnetic confinement  
abs:"tokamak discharge" OR ti:"tokamak discharge"
- soft X-ray diagnostics tokamak plasma  
"soft X-ray" diagnostics tokamak  
abs:"soft x-ray" OR ti:"soft x-ray"

- automatic detection plasma disruptions  
"automatic detection"signals  
abs:"automatic detection"OR ti:"automatic detection"
- machine learning plasma physics fusion  
"machine learning"fusion  
abs:"machine learning"OR ti:"machine learning"
- sawtooth crash soft x-ray tomography  
"sawtooth crash"diagnostics  
abs:"sawtooth crash"OR ti:"sawtooth crash"
- tokamak sawtooth control algorithms  
"sawtooth control"feedback  
abs:"sawtooth control"OR ti:"sawtooth control"
- machine learning automatic detection sawtooth  
"machine learning"sawtooth oscillations  
abs:"sawtooth activity"OR ti:"sawtooth activity"

Затем был произведен подбор в совокупности 12 источников с помощью ответа DeepSeek через поиск в arXiv и первой подборки от chatGPT (начальный список литературы) и сформирован список источников в свободной форме. Этот список был загружен в chatGPT с запросом оформить по ГОСТу. Полученный список был перемещен в текст статьи.

Текст уточняющего промпта:

Оформи выбранные мной источники по госту:

1) Stabilization and Re-excitation of Sawtooth Oscillations due to Energetic Particles in Tokamaks

Authors: H. X. Zhang, H. W. Zhang, Z. W. Ma, J. X. Huang, W. Zhang

2) Comparison of MHD and gyrokinetic simulations of linear instabilities at the  $q = 1$  surface

Authors: F. N. Antlitz, X. Wang, M. Hoelzl, G. T. A. Huijsmans, H. Zhang, J. Puchmayr, Ph. Lauber, T. Hayward-Schneider, B. F. McMillan, A. Mishchenko, E. Poli, Z. X. Lu, JOEKE team

3) A. Gude, M. Maraschek, O. Kardaun et al.

An automated sawtooth detection algorithm for strongly varying plasma conditions and crash characteristics

4) M. Imrisek, V. Weinzettl, J. Mlynar et al.

Use of soft x-ray diagnostic on the COMPASS tokamak for investigations of sawteeth crash neighborhood and of plasma position using fast inversion methods

5) S. von Goeler, W. Stodiek, N. Sauthoff.

Studies of Internal Disruptions and  $m=1$  Oscillations in Tokamak Discharges with Soft-X-Ray Techniques

6) Soft-x-ray tomography of sawtooth oscillations in the JIPP T-II tokama

7) H. X. Zhang, H.W. Zhang et al.

Stabilization and Re-excitation of Sawtooth Oscillations due to Energetic Particles in Tokamaks

8) H. J. OuYang, Z. B. Shi et al.

Real-time identification of sawtooth on HL-3 using a deep learning framework, Plasma Physics and Controlled

9) Francisco Matos, Jakob Svensson, Andrea Pavone et al.

Deep learning for Gaussian process tomography model selection using the ASDEX Upgrade SXR system

10) Detecting Reconnection Events in Kinetic Vlasov Hybrid Simulations Using Clustering Techniques

Authors: Manuela Sisti, Francesco Finelli, Giorgio Pedrazzi, Matteo Faganello, Francesco Califano, Francesca Delli Ponti

11) Neural-Network solver of ideal MHD equilibria

Authors: Timo Thun, Andrea Merlo, Rory Conlin, Dario Panici, Daniel Böckenhoff

12) Automatic ELM detection using gSPRT on the COMPASS tokamak.

Berta, M. Szutyányi, A. Bencze, M. Hron, R. Pánek

13) Automatic identification of edge localized modes in the DIII-D tokamak

Finn H. O'Shea ; Semin Joung ; David R. Smith ; Ryan Coffee

1. Zhang H. X., Zhang H. W., Ma Z. W., Huang J. X., Zhang W.  
Stabilization and re-excitation of sawtooth oscillations due to energetic particles in tokamaks // *Nuclear Fusion*. — 2018. — Vol. 58, № 8. — P. 086028.
2. Antlitz F. N., Wang X., Hoelzl M., Huijsmans G. T. A., Zhang H., Puchmayr J., Lauber Ph., Hayward-Schneider T., McMillan B. F., Mishchenko A., Poli E., Lu Z. X.; JOREK Team.  
Comparison of MHD and gyrokinetic simulations of linear instabilities at the  $q = 1$  surface // *Physics of Plasmas*. — 2020. — Vol. 27, № 4. — P. 042502.
3. Gude A., Maraschek M., Kardaun O., Mlynek A., Zohm H.  
An automated sawtooth detection algorithm for strongly varying plasma conditions and crash characteristics // *Review of Scientific Instruments*. — 2015. — Vol. 86, № 3. — P. 033503.
4. Imrisek M., Weinzettl V., Mlynar J., Melichar J., Bilkova P., Panek R.  
Use of soft X-ray diagnostic on the COMPASS tokamak for investigations of sawteeth crash neighborhood and of plasma position using fast inversion methods // *Review of Scientific Instruments*. — 2014. — Vol. 85, № 11. — P. 11E416.
5. Von Goeler S., Stodiek W., Sauthoff N.  
Studies of internal disruptions and  $m = 1$  oscillations in tokamak discharges with soft X-ray techniques // *Physical Review Letters*. — 1974. — Vol. 33, № 20. — P. 1201–1203.
6. Horiuchi R., Kamimura T., Kawashima H., Miura Y.  
Soft X-ray tomography of sawtooth oscillations in the JIPP T-II tokamak // *Physical Review Letters*. — 1988. — Vol. 61, № 16. — P. 1839–1842.
7. OuYang H. J., Shi Z. B., Liu Y., Wu Y., Xiao B. J.  
Real-time identification of sawtooth on HL-3 using a deep learning framework // *Plasma Physics and Controlled Fusion*. — 2023. — Vol. 65, № 7. — P. 075006.
8. Matos F., Svensson J., Pavone A., Besse N., Guimarães-Filho Z. O.  
Deep learning for Gaussian process tomography model selection using the ASDEX Upgrade soft X-ray system // *Plasma Physics and Controlled Fusion*. — 2020. — Vol. 62, № 10. — P. 105001.
9. Sisti M., Finelli F., Pedrazzi G., Faganello M., Califano F., Delli Ponti F.  
Detecting reconnection events in kinetic Vlasov hybrid simulations using clustering techniques // *Physics of Plasmas*. — 2021. — Vol. 28, № 6. — P. 062305.
10. Thun T., Merlo A., Conlin R., Panici D., Böckenhoff D.  
Neural-network solver of ideal MHD equilibria // *Plasma Physics and Controlled Fusion*. — 2022. — Vol. 64, № 12. — P. 125012.
11. Berta M., Szutyányi M., Bencze A., Hron M., Pánek R.  
Automatic ELM detection using generalized sequential probability ratio test on the COMPASS tokamak // *Plasma Physics and Controlled Fusion*. — 2017. — Vol. 59, № 5. — P. 055001.
12. O'Shea F. H., Joung S., Smith D. R., Coffee R.  
Automatic identification of edge localized modes in the DIII-D tokamak // *Review of Scientific Instruments*. — 2015. — Vol. 86, № 2. — P. 026102.

Рис. 4: Список литературы, оформленный по ГОСТу

С полным протокодом общения с ИИ можно ознакомиться в <source/DeepSeek-arXiv.pdf>

### 2.3.8 Проверка на соответствие структуре научного текста

Последним шагом, когда текст работы уже дополнен диаграммами, видом интерфейса, аннотацией и списком литературы, была произведена проверка на соответствие структуре научного

текста. За основу было взято исследование из лабораторной работы 4, модель chatGPT. Текст промпта:

Вот мой последний вариант текста нир. Проанализируй предоставленный текст на соответствие структуре и стандартам научной работы (статьи или введения к исследованию).

Действуй строго по следующему алгоритму:

1. Идентификация структурных элементов. Найди в тексте и представь в виде таблицы с колонками: «Элемент», «Цитата или констатация отсутствия», «Краткий комментарий». Элементы для поиска: Актуальность исследования, Цель исследования, Задачи исследования, Объект и предмет исследования, Методология / методы, Основные результаты, Выводы

Если элемент отсутствует или выражен неявно – четко отметить это в соответствующей ячейке.

2. Глубинный анализ и оценка.

На основе выявленных элементов проведи критический разбор: Логическая связность: Проследи, вытекают ли задачи из цели, а выводы – из результатов. Оцени, решает ли работа заявленную цель.

Корректность формулировок: Проверь, являются ли цели и задачи конкретными, измеримыми, а предмет и объект – четко разграниченными.

Выявление противоречий: Отметь смысловые или логические противоречия между разделами, а также структурные пробелы (например, отсутствие методологии).

3. Итоговый вердикт.

Представь заключение в трех блоках:

Структурная полнота: Общая оценка наличия всех обязательных элементов научного текста.

Ключевые замечания и ошибки: Список основных проблем, выявленных на этапе глубинного анализа.

Общий вывод: Резюме о соответствии текста структуре научной работы с рекомендациями по устранению главных недостатков.

Важно: Всё время опирайся только на предоставленный текст, не придумывай и не домысливай информацию.

Итоговый вывод от chatGPT (с полным сравнительным анализом и сводной таблицей можно ознакомиться в протоколе общения с ИИ (source/chatGPT.pdf):

*Представленный текст в высокой степени соответствует стандартам научной работы, демонстрирует зрелый уровень постановки задачи, методологии и анализа результатов.*

*Для приведения текста к строго каноническому виду рекомендуется:*

- 1. явно сформулировать объект, предмет и задачи исследования (1–2 абзаца во введении);*
- 2. При необходимости адаптации под статью — сократить описания интерфейса;*
- 3. Структурно разграничить методологию и реализацию.*

*Принципиальных научных недостатков не выявлено — замечания носят формально-структурный характер и легко устранимы.*

*Итоговые протоколы общения с ИИ можно посмотреть в папке source*

### 3 Заключение

Проведённое исследование по автоматизированной генерации научной работы с использованием ИИ позволяет сделать ряд важных выводов. Главный из них состоит в том, что на сегодняшний день **полностью написать полноценную научно-исследовательскую работу или статью исключительно с помощью ИИ не представляется возможным**. Модели способны генерировать внешне связный и стилистически выдержанный текст, но они не могут самостоятельно провести реальное исследование, сформулировать оригинальные научные гипотезы, корректно интерпретировать экспериментальные данные или гарантировать абсолютную достоверность фактов и ссылок. Как показал эксперимент, даже при генерации списка литературы часть предложенных моделями источников оказывалась недостоверной или нерелевантной.

Однако полученные результаты наглядно демонстрируют, что **ИИ может стать мощным инструментом-помощником на различных этапах научной деятельности**. В рамках работы было успешно показано, как с помощью LLM можно:

- **Структурировать и организовать материал:** модели эффективно помогают в создании плана работы, выделении ключевых разделов (актуальность, цель, методы, выводы) и обеспечении логической связности текста.
- **Находить и систематизировать литературу:** при грамотном формулировании промпта ИИ способен предложить тематически релевантные источники (что было подтверждено наличием в сгенерированном списке статей, рекомендованных научным руководителем), а также сформировать поисковые запросы для специализированных архивов, таких как arXiv.
- **Создавать вспомогательные элементы работы:** ИИ успешно справился с генерацией аннотации на русском и английском языках, подбором ключевых слов, созданием UML-диаграммы вариантов использования и даже шаблонов листинга для веб-сервиса.
- **Проводить предварительный структурный анализ текста:** модель может выполнить проверку на наличие обязательных элементов научной работы (цель, задачи, методы, результаты) и указать на формальные недочёты.

Таким образом, современные модели **не заменяют учёного, но существенно экономят его время и ресурсы**, беря на себя рутинные и формализуемые задачи. Они могут выступать в роли интеллектуального ассистента, который помогает исследователю сфокусироваться на сути научной проблемы — постановке задачи, проведении экспериментов, глубоком анализе и интерпретации результатов. Успешное использование ИИ в научной работе требует от человека критического мышления, глубокого понимания предметной области и готовности выступать в роли редактора и валидатора всех сгенерированных материалов.

В перспективе дальнейшее развитие ИИ-инструментов и их интеграция в научный процесс позволят ещё больше автоматизировать подготовительные и сопроводительные этапы исследований, но креативная, аналитическая и ответственная роль исследователя-человека останется незаменимой.